PCT WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6 :
C12N 15/12, C07K 14/705, C12N 15/11, C12Q 1/68, A61K 48/00, C12N 5/10, C07K 16/18, G01N 33/53

A2

(11) International Publication Number: **WO 97/00952**

(43) International Publication Date: 9 January 1997 (09.01.97)

(21) International Application Number: PCT/US96/10618

(22) International Filing Date: 20 June 1996 (20.06.96)

(30) Priority Data:

60/000,352	20 June 1995 (20.06.95)	US
08/567,817	6 December 1995 (06.12.95)	US

(71) Applicant: INCYTE PHARMACEUTICALS, INC. [US/US]; 3174 Porter Drive, Palo Alto, CA 94304 (US).

(72) Inventors: COLEMAN, Roger; 260 Mariposa #2, Mountain View, CA 94041 (US). GUEGLER, Karl, J.; 1048 Oakland Avenue, Menlo Park, CA 94025 (US). AU-YOUNG, Janice; 1419 Kains Avenue, Berkeley, CA 94702 (US). BANDMAN, Olga; 2309 Rock Street #27, Mountain View, CA 94043 (US). SEILHAMER, Jeffrey, J.; 12555 La Cresta, Los Altos Hills, CA 94022 (US).

(74) Agent: GLAISTER, Debra, J.; Incyte Pharmaceuticals, Inc., 3174 Porter Drive, Palo Alto, CA 94304 (US).

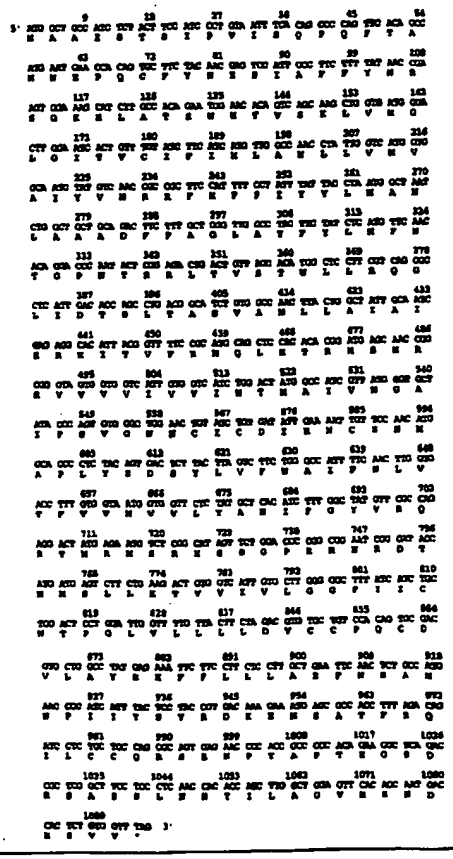
(81) Designated States: AT, AU, BR, CA, CH, CN, DE, DK, ES, FI, GB, IL, JP, KR, MX, NO, NZ, RU, SE, SG, ARIPO patent (KE, LS, MW, SD, SZ, UG), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, ML, MR, NE, SN, TD, TG).

Published
Without international search report and to be republished upon receipt of that report.

(54) Title: A HUMAN EDG-2 RECEPTOR HOMOLOG

(57) Abstract

The present invention provides nucleic acid and amino acid sequences that identify and encode a novel EDG-2 receptor homolog (hedg) expressed in human rheumatoid synovium. The present invention also provides for probes for the detection of nucleotide sequences encoding of HEDG or HEDG-like molecules, antisense molecules to the nucleotide sequences which encode HEDG, diagnostic tests based on HEDG encoding nucleic acid molecules, genetically engineered expression vectors and host cells for the production of purified HEDG, antibodies capable of binding specifically to HEDG, and antagonists and inhibitors with specific binding activity for the polypeptide HEDG.



FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AM	Armenia	**GB**	United Kingdom	**MW**	Malawi
AT	Austria	**GE**	Georgia	**MX**	Mexico
AU	Australia	**GN**	Guinea	**NE**	Niger
BB	Barbados	**GR**	Greece	**NL**	Netherlands
BE	Belgium	**HU**	Hungary	**NO**	Norway
BF	Burkina Faso	**IE**	Ireland	**NZ**	New Zealand
BG	Bulgaria	**IT**	Italy	**PL**	Poland
BJ	Benin	**JP**	Japan	**PT**	Portugal
BR	Brazil	**KE**	Kenya	**RO**	Romania
BY	Belarus	**KG**	Kyrgystan	**RU**	Russian Federation
CA	Canada	**KP**	Democratic People's Republic of Korea	**SD**	Sudan
CF	Central African Republic			**SE**	Sweden
CG	Congo	**KR**	Republic of Korea	**SG**	Singapore
CH	Switzerland	**KZ**	Kazakhstan	**SI**	Slovenia
CI	Côte d'Ivoire	**LI**	Liechtenstein	**SK**	Slovakia
CM	Cameroon	**LK**	Sri Lanka	**SN**	Senegal
CN	China	**LR**	Liberia	**SZ**	Swaziland
CS	Czechoslovakia	**LT**	Lithuania	**TD**	Chad
CZ	Czech Republic	**LU**	Luxembourg	**TG**	Togo
DE	Germany	**LV**	Latvia	**TJ**	Tajikistan
DK	Denmark	**MC**	Monaco	**TT**	Trinidad and Tobago
EE	Estonia	**MD**	Republic of Moldova	**UA**	Ukraine
ES	Spain	**MG**	Madagascar	**UG**	Uganda
FI	Finland	**ML**	Mali	**US**	United States of America
FR	France	**MN**	Mongolia	**UZ**	Uzbekistan
GA	Gabon	**MR**	Mauritania	**VN**	Viet Nam

A HUMAN EDG-2 RECEPTOR HOMOLOG

TECHNICAL FIELD

The present invention is in the field of molecular biology; more particularly, the present invention describes a nucleic acid sequence and an amino acid sequence for a novel human EDG-2 receptor homolog.

BACKGROUND ART

The EDG-2 receptor is a putative G-protein coupled seven transmembrane receptor (T7G) which was initially cloned from sheep mRNA (GenBank U18405; Masana MI et al (1994) unpublished). Human edg-1 is commonly grouped with orphan receptors because its endogenous ligand is not known (Hla T and Maciag T (1990) J Biol Chem 265:9308-13). Several T7G receptors have been classified as orphan receptors; they include LCR-1 from brain, the mas oncogene associated with epidermoid carcinoma, RDC-1 known from several major organs, and R334 from rat brain and testis. In some of these cases, a ligand was initially proposed and has since been discounted. The orphan receptors vary in number of amino acids, in molecular weight, in glycosylation sites, and presence and number of disulfide bonds (Watson S and Arkinstall S (1994) The G-Protein Linked Receptor Facts Book, Academic Press, San Diego CA).

They are, however, related to other T7Gs by their seven hydrophobic domains which span the plasma membrane and form a bundle of antiparallel α helices. These transmembrane segments (TMS) are designated by roman numerals I-VII and account for structural and functional features of the receptor. In most cases, the bundle of helices forms a binding pocket; however, when the binding site must accommodate more bulky molecules, the extracellular N-terminal segment or one or more of the three extracellular loops participate in binding and in subsequent induction of conformational change in intracellular portions of the receptor. The activated receptor, in turn, interacts with an intracellular G-protein complex which mediates further intracellular signalling activities generally the production of second messengers such as cyclic AMP (cAMP), phospholipase C, inositol triphosphate or ion channel proteins.

T7G receptors are expressed and activated during numerous developmental and disease processes. Identification of a novel T7G receptor provides the opportunity to diagnose or intervene in such processes, and the receptor can be used in screening assays to identify physiological or pharmaceutical molecules which trigger, prolong or inhibit its activity.

DISCLOSURE OF THE INVENTION

The subject invention provides a unique nucleotide sequence which encodes a novel human EDG-2 receptor homolog (HEDG). The cDNA, herein designated hedg, was identified and cloned using Incyte Clone No. 80853 from a rheumatoid synovium cDNA library.

The invention relates to the use of nucleic acid and amino acid sequences of HEDG, or its variants, in the diagnosis or treatm nt of activated, inflamed or diseas d cells and/or tissues associated with its expression. Aspects of th invention includ the antisens DNA of hedg; cloning or expression vectors containing hedg; host cells or organisms transformed with expression vectors containing hedg; a method for the production and recovery of purified HEDG from host cells; and purified protein, HEDG, which can be used to identify inhibitors for the downregulation of signal transduction involving HEDG.

BRIEF DESCRIPTION OF DRAWINGS

Figures 1A and 1B shows the alignment of the nucleic acid sequence (coding region of SEQ ID NO: 1) and amino acid sequence (SEQ ID NO:2) for HEDG. The alignment of the sequences was produced using MacDNAsis software (Hitachi Software Engineering Co Ltd)

Figure 2 displays the alignment of HEDG with sheep EDG-2 (U18405; SEQ ID NO:3) and human EDG-1 (GI 119130; SEQ ID NO:4) receptors. Note the conserved Arg^{36} and Ser^{37} cleavage site characteristic of these T7G molecules. Sequences for Fig. 2 were aligned using the multisequence alignment program of DNAStar software (DNAStar Inc, Madison WI).

MODES FOR CARRYING OUT THE INVENTION

As used herein and designated by the upper case abbreviation, HEDG, refers to an EDG2 receptor homolog in either naturally occurring or synthetic form and active fragments thereof which have the amino acid sequence of SEQ ID NO:2. In one embodiment, the polypeptide HEDG is encoded by mRNAs transcribed from the cDNA, as designated by the lower case abbreviation, hedg, of SEQ ID NO:1.

The novel human edg-2 receptor homolog, HEDG, which is the subject of this patent application, was discovered among the partial cDNA sequences (Incyte Clone 80853) expressed in a rheumatoid synovium library. It is more distantly homologous to human edg-1 which was cloned from human vascular endothelial cells and expressed in epithelioid cells, fibroblasts, melanocytes, and vascular smooth muscle cells.

An "oligonucleotide" is a stretch of nucleotide residues which has a sufficient number of bases to be used as an oligomer, amplimer or probe in a polymerase chain reaction (PCR). Oligonucleotides are prepared from genomic or cDNA sequence and are used to amplify, reveal or confirm the presence of a similar DNA or RNA in a particular cell or tissue. Oligonucleotides or oligomers comprise portions of a DNA sequence having at least about 10 nucleotides and as many as about 35 nucleotides, preferably about 25 nucleotides.

"Probes" may be derived from naturally occurring or recombinant single- or double-stranded nucleic acids or be chemically synthesized. They are useful in detecting the presence of identical or similar sequences.

A "portion" or "fragment" of a polynucleotide or nucleic acid compris s all or any part of the nucleotid sequence having fewer nucleotides than about 6 kb, preferably few r than about 1 kb which can be used as a probe. Such probes may be labelled with reporter molecules using, nick translation, Klenow fill-in reaction, PCR or other methods well known in the art. After pretesting to optimize reaction conditions and to eliminate false positives, nucleic acid probes may be used in Southern, northern or in situ hybridizations to determine whether DNA or RNA encoding HEDG is present in a cell type, tissue, or organ.

"Reporter" molecules are those radionuclides, enzymes, fluorescent, chemiluminescent, or chromogenic agents which associate with, establish the presence of, and may allow quantification of a particular nucleotide or amino acid sequence.

"Recombinant nucleotide variants" encoding HEDG may be synthesized by making use of the "redundancy" in the genetic code. Various codon substitutions, such as the silent changes which produce specific restriction sites or codon usage-specific mutations, may be introduced to optimize cloning into a plasmid or viral vector or expression in a particular prokaryotic or eukaryotic host system, respectively.

"Chimeric" molecules may be constructed by introducing all or part of the nucleotide sequence of this invention into a vector containing additional nucleic acid sequence which might be expected to change any one (or more than one) of the following HEDG characteristics: cellular location, distribution, ligand-binding affinities, interchain affinities, degradation/turnover rate, signalling, etc.

"Active" refers to those forms, fragments, or domains of any HEDG polypeptide which retain the biologic and/or antigenic activities of any naturally occurring HEDG.

"Naturally occurring HEDG" refers to a polypeptide produced by cells which have not been genetically engineered and specifically contemplates various polypeptides arising from post-translational modifications of the polypeptide including but not limited to acetylation, carboxylation, glycosylation, phosphorylation, lipidation and acylation.

"Derivative" refers to those polypeptides which have been chemically modified by such techniques as ubiquitination, labelling (see above), pegylation (derivatization with polyethylene glycol), and chemical insertion or substitution of amino acids such as ornithine which do not normally occur in human proteins.

"Recombinant polypeptide variant" refers to any polypeptide which differs from naturally occurring HEDG by amino acid insertions, deletions and/or substitutions, created using recombinant DNA techniques. Guidance in determining which amino acid residues may be replaced, added or deleted without abolishing activities of interest may be found by comparing the sequence of HEDG with that of related polypeptides and minimizing the number of amino acid

sequence changes made in highly conserv d regions.

Amino acid "substitutions" are conservative in nature when they result from replacing one amino acid with another having similar structural and/or chemical properties, such as the replacement of a leucine with an isoleucine or valine, an aspartate with a glutamate, or a threonine with a serine.

"Insertions" or "deletions" are typically in the range of about 1 to 5 amino acids. The variation allowed may be experimentally determined by producing the peptide synthetically or by systematically making insertions, deletions, or substitutions of nucleotides in the hedg sequence using recombinant DNA techniques.

A "signal or leader sequence" can be used, when desired, to direct the polypeptide through a membrane of a cell. Such a sequence may be naturally present on the polypeptides of the present invention or provided from heterologous sources by recombinant DNA techniques.

An "oligopeptide" is a short stretch of amino acid residues and may be expressed from an oligonucleotide. It may be functionally equivalent to and the same length as (or considerably shorter than) a "fragment," "portion," or "segment" of a polypeptide. Such sequences comprise a stretch of amino acid residues of at least about 5 amino acids and often about 17 or more amino acids, typically at least about 9 to 13 amino acids, and of sufficient length to display biologic and/or antigenic activity.

"Inhibitor" is any substance which retards or prevents a chemical or physiological reaction or response. Common inhibitors include but are not limited to antisense molecules, antibodies, and antagonists.

"Standard" expression is a quantitative or qualitative measurement for comparison. It is based on a statistically appropriate number of normal samples and is created to use as a basis of comparison when performing diagnostic assays, running clinical trials, or following patient treatment profiles.

"Animal" as used herein may be defined to include human, domestic (cats dogs, etc.), agricultural (cows, horses, sheep, etc) or test species (mouse, rat, rabbit, etc).

The present invention provides a nucleotide sequence uniquely identifying a novel seven transmembrane receptor, human EDG-2 or HEDG. Because HEDG is specifically expressed in inflamed rheumatoid synovium, the nucleic acids (hedg), polypeptides (HEDG) and antibodies to HEDG are useful in diagnostic assays which survey for increased receptor production. Excessive expression of HEDG is likely to be associated with the activation of T lymphocytes and other cells which respond to inflammation and can result in the production of abundant proteases and other molecules which can lead to tissue damage or destruction. Therefore, a diagnostic test for excessive expression of HEDG can accelerate diagnosis and proper treatment

of abnormal conditions caused by viral, bacterial or fungal infections; allergic responses; mechanical injury associated with trauma; hereditary diseases; lymphoma or carcinoma; or oth r conditions which activate the genes of lymphoid tissues.

The nucleotide sequences encoding HEDG (or their complement) have numerous applications in techniques known to those skilled in the art of molecular biology. These techniques include use as hybridization probes, use in the construction of oligomers for PCR, use for chromosome and gene mapping, use in the recombinant production of HEDG, and use in generation of antisense DNA or RNA, their chemical analogs and the like. Uses of nucleotides encoding HEDG disclosed herein are exemplary of known techniques and are not intended to limit their use in any technique known to a person of ordinary skill in the art. Furthermore, the nucleotide sequences disclosed herein may be used in molecular biology techniques that have not yet been developed, provided the new techniques rely on properties of nucleotide sequences that are currently known, e.g., the triplet genetic code, specific base pair interactions, etc.

It will be appreciated by those skilled in the art that as a result of the degeneracy of the genetic code, a multitude of HEDG-encoding nucleotide sequences may be produced. Some of these will only bear minimal homology to the nucleotide sequence of the known and naturally occurring HEDG. The invention has specifically contemplated each and every possible variation of nucleotide sequence that could be made by selecting combinations based on possible codon choices. These combinations are made in accordance with the standard triplet genetic code as applied to the nucleotide sequence of naturally occurring hedg, and all such variations are to be considered as being specifically disclosed.

Although the nucleotide sequences which encode HEDG, its derivatives or its variants are preferably capable of hybridizing to the nucleotide sequence of the naturally occurring hedg under stringent conditions, it may be advantageous to produce nucleotide sequences encoding HEDG or its derivatives possessing a substantially different codon usage. Codons can be selected to increase the rate at which expression of the peptide occurs in a particular prokaryotic or eukaryotic expression host in accordance with the frequency with which particular codons are utilized by the host. Other reasons for substantially altering the nucleotide sequence encoding HEDG and/or its derivatives without altering the encoded aa sequence include the production of RNA transcripts having more desirable properties, such as a greater half-life, than transcripts produced from the naturally occurring sequence.

Nucleotide sequences encoding HEDG may be joined to a variety of other nucleotide sequences by means of well established recombinant DNA techniques (Sambrook J et al (1989) Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Laboratory, Cold Spring Harbor NY; or Ausubel FM et al (1989) Current Protocols in Molecular Biology, John Wiley & Sons,

New York City). Useful nucl otide sequences for joining to hedg include an assortment of cloning vectors such as plasmids, cosmids, lambda phage derivatives, phagemids, and the like. Vectors of interest include expression vectors, replication vectors, probe generation vectors, sequencing vectors, etc. In general, vectors of interest may contain an origin of replication functional in at least one organism, convenient restriction endonuclease sensitive sites, and selectable markers for one or more host cell systems.

Another aspect of the subject invention is to provide for hedg-specific hybridization probes capable of hybridizing with naturally occurring nucleotide sequences encoding HEDG. Such probes may also be used for the detection of similar T7G encoding sequences and should preferably contain at least 50% of the nucleotides from the hedg sequence. The hybridization probes of the subject invention may be derived from the nucleotide sequence presented as SEQ ID NO:1 or from genomic sequences including promoter, enhancers or introns of the native gene. Hybridization probes may be labeled by a variety of reporter molecules using techniques well known in the art.

PCR as described US Patent Nos. 4,683,195; 4,800,195; and 4,965,188 provides additional uses for oligonucleotides based upon the nucleotide sequence which encodes HEDG. Such probes used in PCR may be of recombinant origin, chemically synthesized, or a mixture of both. Oligomers may comprise discrete nucleotide sequences employed under optimized conditions for identification of hedg in specific tissues or diagnostic use. The same two oligomers, a nested set of oligomers, or even a degenerate pool of oligomers may be employed under less stringent conditions for identification of closely related DNAs or RNAs.

Other means of producing specific hybridization probes for hedg include the cloning of nucleic acid sequences encoding HEDG or HEDG derivatives into vectors for the production of mRNA probes. Such vectors are known in the art, are commercially available and may be used to synthesize RNA probes in vitro by means of the addition of the appropriate RNA polymerase as T7 or SP6 RNA polymerase and the appropriate reporter molecules.

It is possible to produce a DNA sequence, or portions thereof, entirely by synthetic chemistry. After synthesis, the nucleic acid sequence can be inserted into any of the many available DNA vectors and their respective host cells using techniques which are well known in the art. Moreover, synthetic chemistry may be used to introduce mutations into the nucleotide sequence. Alternately, a portion of sequence in which a mutation is desired can be synthesized and recombined with longer portion of an existing genomic or recombinant sequence.

The nucleotide sequence for hedg can be used in an assay to detect inflammation or disease associated with abnormal levels of HEDG expression. The cDNA can be labeled by methods known in the art, added to a fluid, cell or tissue sample from a patient, and incubated under hybridizing

conditions. After an incubation period, the sample is washed with a compatible fluid which optionally contains a reporter molecule. After the compatible fluid is rinsed off, the reporter molecule is quantitated and compared with a standard as previously defined. If kinase expression is significantly different from standard expression, the assay indicates inflammation or disease.

The nucleotide sequence for hedg can be used to construct hybridization probes for mapping the native gene. The gene may be mapped to a particular chromosome or to a specific region of a chromosome using well known mapping techniques. These techniques include in situ hybridization of chromosomal spreads (Verma et al (1988) Human Chromosomes: A Manual of Basic Techniques, Pergamon Press, New York City), flow-sorted chromosomal preparations, or artificial chromosome constructions such as yeast artificial chromosomes (YACs), bacterial artificial chromosomes (BACs), bacterial P1 constructions or single chromosome cDNA libraries.

In situ hybridization of chromosomal preparations and physical mapping techniques such as linkage analysis using established chromosomal markers are invaluable in extending genetic maps. Examples of genetic map data can be found in the yearly genome issue of Science (eg 1994, 265:1981f). Often locating a gene on the chromosome of another mammalian species may reveal associated markers which can be used to help identify the analogous human chromosome.

New nucleotide sequences can be assigned to chromosomal subregions by physical mapping. The mapping of new genes or nucleotide sequences provide useful landmarks for investigators searching for disease genes using positional cloning or other gene discovery techniques. Once a disease or syndrome, such as ataxia telangiectasia (AT), has been crudely localized by genetic linkage to a particular genomic region, for example, AT to 11q22-23 (Gatti et al (1988) Nature 336:577-580), any sequences mapping to that area may represent or reveal genes for further investigation. The nucleotide sequence of the subject invention may also be used to detect differences in gene sequence between normal and carrier or affected individuals.

Nucleotide sequences encoding hedg may be used to produce a purified oligo- or polypeptide using well known methods of recombinant DNA technology. Goeddel (1990, Gene Expression Technology, Methods and Enzymology, Vol 185, Academic Press, San Diego CA) is one among many publications which teach expression of an isolated nucleotide sequence. The oligopeptide may be expressed in a variety of host cells, either prokaryotic or eukaryotic. Host cells may be from the same species from which the nucleotide sequence was derived or from a different species. Advantages of producing an oligonucleotide by recombinant DNA technology

include obtaining adequat amounts of the protein for purification and the availability of simplified purification procedures.

Cells transformed with DNA encoding HEDG may be cultured under conditions suitable for the expression of T7Gs, their extracellular, transmembrane or intracellular domains and recovery of such peptides from cell culture. HEDG (or any of its domains) produced by a recombinant cell may be secreted or may be contained intracellularly, depending on the particular genetic construction used. In general, it is more convenient to prepare recombinant proteins in secreted form. Purification steps vary with the production process and the particular protein produced. Often an oligopeptide can be produced from a chimeric nucleotide sequence. This is accomplished by ligating the nucleotides from hedg or a desired portion of the polypeptide to a nucleic acid sequence encoding a polypeptide domain which will facilitate protein purification (Kroll DJ et al (1993) DNA Cell Biol 12:441-53).

In addition to recombinant production, fragments of HEDG may be produced by direct peptide synthesis using solid-phase techniques (eg Stewart et al (1969) Solid-Phase Peptide Synthesis, WH Freeman Co, San Francisco CA; Merrifield J (1963) J Am Chem Soc 85:2149-2154. Automated synthesis may be achieved, for example, using Applied Biosystems 431A Peptide Synthesizer (Foster City, CA) in accordance with the instructions provided by the manufacturer. Additionally, a particular portion of HEDG may be mutated during direct synthesis and combined with other parts of the peptide using chemical methods.

HEDG for antibody induction does not require biological activity; however, the protein must be antigenic. Peptides used to induce specific antibodies may have an aa sequence consisting of at least five aa, preferably at least 10 aa. They should mimic a portion of the aa sequence of the protein and may contain the entire aa sequence of a small naturally occurring molecule such as HEDG. An antigenic portion of HEDG may be fused to another protein such as keyhole limpet hemocyanin, and the chimeric molecule used for antibody production.

Antibodies specific for HEDG may be produced by inoculation of an appropriate animal with the polypeptide or an antigenic fragment. An antibody is specific for HEDG if it is produced against an epitope of the polypeptide and binds to at least part of the natural or recombinant protein. Antibody production includes not only the stimulation of an immune response by injection into animals, but also analogous processes such as the production of synthetic antibodies, the screening of recombinant immunoglobulin libraries for specific-binding molecules (eg Orlandi R et al (1989) PNAS 86:3833-3837, or Huse WD et al (1989) Science 256:1275-1281) or the in vitro stimulation of lymphocyte populations. Current technology (Winter G and Milstein C (1991) Nature 349:293-299) provides for a number of highly specific binding reagents based on the principles of antibody formation. These techniques

may be adapted to produce molecules which specifically bind HEDGs.

An additional embodiment of the subject invention is the use of HEDG specific antibodies, inhibitors, receptors or their analogs as bioactive agents to treat inflammation or disease including, but not limited to viral, bacterial or fungal infections; allergic responses; mechanical injury associated with trauma; hereditary diseases; lymphoma or carcinoma; or other conditions which activate the genes of lymphoid tissues.

Bioactive compositions comprising agonists, antagonists, receptors or inhibitors of HEDG may be administered in a suitable therapeutic dose determined by any of several methodologies including clinical studies on mammalian species to determine maximal tolerable dose and on normal human subjects to determine safe dose. Additionally, the bioactive agent may be complexed with a variety of well established compounds or compositions which enhance stability or pharmacological properties such as half-life. It is contemplated that the therapeutic, bioactive composition may be delivered by intravenous infusion into the bloodstream or any other effective means which could be used for treating problems involving excessive lymphocyte and leukocyte trafficking.

Rheumatoid arthritis is currently evaluated on the basis of swelling, response to NSAIDs, x-rays, etc. HEDG is most likely expressed on the surface of the fibroblasts, T and B lymphocytes, monocyte/macrophages, or mast cells which comprise the cells of the inflamed synovium. Once adequate standards are established, an assay for the abnormal expression of HEDG is a viable diagnostic tool for assessing the extent that RA has progressed. The expression of HEDG in a sustained inflammatory response makes it a valuable therapeutic target for screening drug libraries. Inhibitors of HEDG are useful for controlling signal transduction and signaling cascades in cells of the rheumatoid synovium.

The examples below are provided to illustrate the subject invention. These examples are provided by way of illustration and are not included for the purpose of limiting the invention.

INDUSTRIAL APPLICABILITY

I Isolation of mRNA and Construction of the cDNA Library

The hedg sequence of this application was first identified in Incyte Clone No. 80853 among the sequences comprising the rheumatoid synovium library. Rheumatoid synovial tissue was obtained from the hip joint removed from a 68 year old female with erosive, nodular rheumatoid arthritis. The tissue was frozen, ground to powder in a mortar and pestle, and lysed immediately in buffer containing guanidinium isothiocyanate. Lysis was followed by several phenol-chloroform extractions and ethanol precipitations. Poly-A+ mRNA was isolated using biotinylated oligo d(T) and streptavidin coupled to paramagnetic particles (Poly(A) Tract Isolation System, Promega, Madison WI).

Using this Poly-A+ mRNA, a custom cDNA library was constructed by Stratagen (La Jolla CA). Synthesis of cDNA was primed with oligo d(T), and adapter oligonucleotides were ligated onto the cDNA molecules enabling th m to be inserted into the Uni-ZAP™ vector system (Stratagene). Alternative unidirectional vectors might include, but are not limited to, pcDNAI (Invitrogen, San Diego CA) and pSHlox-1 (Novagen, Madison WI).

I I Isolation of cDNA Clones

The phagemid forms of individual cDNA clones were obtained by the in vivo excision process, in which the host bacterial strain was co-infected with both the library phage and an f1 helper phage. Polypeptides or enzymes derived from both the library-containing phage and the helper phage nicked the DNA, initiated new DNA synthesis from defined sequences on the target DNA, and created a smaller, single stranded circular phagemid DNA molecule that included all DNA sequences of the pBluescript phagemid and the cDNA insert. The phagemid DNA was released from the cells, purified, and used to reinfect fresh host cells (SOLR™, Stratagene) where double-stranded DNA was produced.

DNA was purified using the QIAWELL-8 Plasmid Purification System (QIAGEN Inc, Chatsworth CA) an anion-exchange resin system with EMPORE™ membrane technology (3M, Minneapolis MN). The DNA was eluted from the purification resin and prepared for DNA sequencing and other analytical manipulations.

III Sequencing of cDNA Clones

The cDNA inserts from random isolates of the rheumatoid synovium library were sequenced in part. Methods for DNA sequencing are well known in the art. Conventional enzymatic methods employed DNA polymerase Klenow fragment, SEQUENASE® (US Biochemical Corp, Cleveland OH) or Taq polymerase to extend DNA chains from an oligonucleotide primer annealed to the DNA template of interest. Methods have been developed for the use of both single- and double-stranded templates. The chain termination reaction products were electrophoresed on urea-acrylamide gels and detected either by autoradiography (for radionuclide-labeled precursors) or by fluorescence (for fluorescent-labeled precursors). Recent improvements in mechanized reaction preparation, sequencing and analysis using the fluorescent detection method have permitted expansion in the number of sequences that are determined per day using machines such as the Catalyst 800 and the Applied Biosystems 377 or 373 DNA sequencers.

V Homology Searching of cDNA Clones and Deduced Proteins

Each sequence so obtained was compared to sequences in GenBank using a search algorithm developed by Applied Biosystems and incorporated into the INHERIT™ 670 Sequence Analysis System. In this algorithm, Pattern Specification Language (developed by TRW Inc, Los

Angeles CA) was used to determine regions of homology. The three parameters that determine how the sequence comparisons run were window size, window offset, and error tolerance. Using a combination of these three parameters, the DNA database was searched for sequences containing regions of homology to the query sequence, and the appropriate sequences were scored with an initial value. Subsequently, these homologous regions were examined using dot matrix homology plots to distinguish regions of homology from chance matches. Smith-Waterman alignments were used to display the results of the homology search.

Peptide and protein sequence homologies were ascertained using the INHERIT™ 670 Sequence Analysis System in a way similar to that used in DNA sequence homologies. Pattern Specification Language and parameter windows were used to search protein databases for sequences containing regions of homology which were scored with an initial value. Dot-matrix homology plots were examined to distinguish regions of significant homology from chance matches.

Alternatively, BLAST, which stands for Basic Local Alignment Search Tool, is used to search for local sequence alignments (Altschul SF (1993) J Mol Evol 36:290-300; Altschul, SF et al (1990) J Mol Biol 215:403-10). BLAST produces alignments of both nucleotide and aa sequences to determine sequence similarity. Because of the local nature of the alignments, BLAST is especially useful in determining exact matches or in identifying homologs. Whereas it is ideal for matches which do not contain gaps, it is inappropriate for performing motif-style searching. The fundamental unit of BLAST algorithm output is the High-scoring Segment Pair (HSP).

An HSP consists of two sequence fragments of arbitrary but equal lengths whose alignment is locally maximal and for which the alignment score meets or exceeds a threshold or cutoff score set by the user. The BLAST approach is to look for HSPs between a query sequence and a database sequence, to evaluate the statistical significance of any matches found, and to report only those matches which satisfy the user-selected threshold of significance. The parameter E establishes the statistically significant threshold for reporting database sequence matches. E is interpreted as the upper bound of the expected frequency of chance occurrence of an HSP (or set of HSPs) within the context of the entire database search. Any database sequence whose match satisfies E is reported in the program output.

V Identification, Full Length Cloning, Sequencing and Translation

Analysis of INHERIT™ results from randomly picked and sequenced portions of clones from the rheumatoid synovium library identified Incyte 80853 as a homolog of sheep EDG-2 receptor. The partial sequence displayed 92.6% identity with nucleotide sequence of accession U18405 in GenBank (Masana MI et al, supra). The cDNA insert comprising Incyte 80853 was

fully sequenced and used as th basis for cloning the full length cDNA.

The cDNA was extended to full length using a modified XL-PCR (Perkin Elmer) procedure as disclosed in Patent Application Serial No 08/487,112 filed 7 June 1995, and the rheumatoid synovium cDNA library as a template. Primers were designed based on known sequence; one primer was synthesized to initiate extension in the antisense direction (XLR = TCATCTTGATCCCCATCCCTTCTG) and the other to extend sequence in the sense direction (XLF = AGTCTCCGAGTATTGGGTCCTGTG). The primers allowed the sequence to be extended "outward" generating amplicons containing new, unknown nucleotide sequence for the gene of interest. The primers were designed using Oligo 4.0 (National Biosciences Inc, Plymouth MN) to be 22-30 nucleotides in length, to have a GC content of 50% or more, and to anneal to the target sequence at temperatures about 68°-72° C. Any stretch of nucleotides which would result in hairpin structures and primer-primer dimerizations were avoided.

By following the instructions for the XL-PCR kit and thoroughly mixing the enzyme and reaction mix, high fidelity amplification is obtained. Beginning with 25 pMol of each primer and the recommended concentrations of all other components of the kit, PCR was performed using the Peltier Thermal Cycler (PTC200; MJ Research, Watertown MA) and the following parameters:

Step 1	94° C for 60 sec (initial denaturation)
Step 2	94° C for 15 sec
Step 3	65° C for 1 min
Step 4	68° C for 7 min
Step 5	Repeat step 2-4 for 15 additional cycles
Step 6	94° C for 15 sec
Step 7	65° C for 1 min
Step 8	68° C for 7 min + 15 sec/cycle
Step 9	Repeat step 6-8 for 11 additional cycles
Step 10	72° C for 8 min
Step 11	4° C (and holding)

At the end of 28 cycles, 50 µl of the reaction mix was removed; and the remaining reaction mix was run for an additional 10 cycles as outlined below:

Step 1	94° C for 15 sec
Step 2	65° C for 1 min
Step 3	68° C for (10 min + 15 sec)/cycle
Step 4	Repeat step 1-3 for 9 additional cycles
Step 5	72° C for 10 min

A 5-10 µl aliquot of the reaction mixture was analyzed by electrophoresis on a low concentration (about 0.6-0.8%) agarose mini-gel to determine which reactions were successful in extending the sequence. Although all extensions potentially contain a full length gene, some of the largest products or bands were selected and cut out of the gel. Further purification involved using a commercial gel extraction method such as QIAQuick™ (QIAGEN

Inc). After recovery of the DNA, Klenow enzyme was used to trim single-stranded, nucleotide overhangs creating blunt ends which facilitated religation and cloning.

After ethanol precipitation, the products w r r dissolved in 13 µl of ligation buffer. Then, 1µl T4-DNA ligase (15 units) and 1µl T4 polynucleotide kinase were added, and the mixture was incubated at room temperature for 2-3 hours or overnight at 16° C. Competent E. coli cells (in 40 µl of appropriate media) were transformed with 3 µl of ligation mixture and cultured in 80 µl of SOC medium (Sambrook J et al, supra). After incubation for one hour at 37° C, the whole transformation mixture was plated on Luria Bertani (LB)-agar (Sambrook J et al, supra) containing carbenicillin at 25 mg/L (2xCarb). The following day, 12 colonies were randomly picked from each plate and cultured in 150 µl of liquid LB/2xCarb medium placed in an individual well of an appropriate, commercially-available, sterile 96-well microtiter plate. The following day, 5 µl of each overnight culture was transferred into a non-sterile 96-well plate and after dilution 1:10 with water, 5 µl of each sample was transferred into a PCR array.

For PCR amplification, 15 µl of concentrated PCR reaction mix (1.33X) containing 0.75 units of Taq polymerase, a vector primer and one or both of the gene specific primers used for the extension reaction were added to each well. Amplification was performed using the following conditions:

Step 1	94° C for 60 sec
Step 2	94° C for 20 sec
Step 3	55° C for 30 sec
Step 4	72° C for 90 sec
Step 5	Repeat steps 2-4 for an additional 29 cycles
Step 6	72° C for 180 sec
Step 7	4° C (and holding)

Aliquots of the PCR reactions were run on agarose gels together with molecular weight markers. The sizes of the PCR products were compared to the original partial cDNAs, and appropriate clones were selected, ligated into plasmid and sequenced.

The nucleotide and aa sequences for human HEDG are shown in Figure 1. The coding region of hedg begins at nucleotide 309 and ends at nucleotide 1403 of SEQ ID NO:1. When the three possible translations of HEDG were searched against protein databases such as SwissProt and PIR, no exact matches were found. Figure 2 shows the comparison between the amino acid sequences of HEDG, sheep EDG-2 (U18405) and human EDG-1 (GI 119130).

VI Antisense analysis

Knowledge of the correct, complete cDNA sequence of HEDG enables its use as a tool for antisense technology in the investigation of gene function. Oligonucleotides, cDNA or genomic fragments comprising the antisense strand of hedg are used either in vitro or in vivo to inhibit expression of the mRNA. Such technology is now well known in the art, and antisense molecules

can be designed at various locations along the nucleotide sequ nces. By treatment of cells or whole test animals with such antisense sequences, the gen of interest is ffectively turned off. Frequently, the function of the gen is ascertained by observing behavior at the intracellular, cellular, tissue or organismal level (eg, lethality, loss of differentiated function, changes in morphology, etc).

In addition to using sequences constructed to interrupt transcription of a particular open reading frame, modifications of gene expression is obtained by designing antisense sequences to intron regions, promoter/enhancer elements, or even to trans-acting regulatory genes. Similarly, inhibition is achieved using Hogeboom base-pairing methodology, also known as "triple helix" base pairing.

VII Expression of HEDG

Expression of hedg is accomplished by subcloning the cDNAs into appropriate expression vectors and transfecting the vectors into analogous expression hosts. In this particular case, the cloning vector previously used for the generation of the cDNA library also provides for direct expression of hedg sequences in E. coli. Upstream of the cloning site, this vector contains a promoter for ß-galactosidase, followed by sequence containing the amino-terminal Met and the subsequent 7 residues of ß-galactosidase. Immediately following these eight residues is an engineered bacteriophage promoter useful for artificial priming and transcription and for providing a number of unique endonuclease restriction sites for cloning.

Induction of the isolated, transfected bacterial strain with IPTG using standard methods produces a fusion protein corresponding to the first seven residues of ß-galactosidase, about 15 residues of "linker", and the peptide encoded within the cDNA. Since cDNA clone inserts are generated by an essentially random process, there is one chance in three that the included cDNA will lie in the correct frame for proper translation. If the cDNA is not in the proper reading frame, it is obtained by deletion or insertion of the appropriate number of bases using well known methods including in vitro mutagenesis, digestion with exonuclease III or mung bean nuclease, or the inclusion of an oligonucleotide linker of appropriate length.

The hedg cDNA is shuttled into other vectors known to be useful for expression of protein in specific hosts. Oligonucleotide primers containing cloning sites as well as a segment of DNA (about 25 bases) sufficient to hybridize to stretches at both ends of the target cDNA is synthesized chemically by standard methods. These primers are then used to amplify the desired gene segment by PCR. The resulting gene segment is digested with appropriate restriction enzymes under standard conditions and isolated by gel electrophoresis. Alternately, similar gene segments are produced by digestion of the cDNA with appropriate restriction enzymes. Using appropriate primers, segments of coding sequence from more than one gene are

ligated together and cloned in appropriate vectors. It is possible to optimize expression by construction of such chimeric sequences.

Suitable expression hosts for such chimeric molecules include, but are not limited to, mammalian cells such as Chinese Hamster Ovary (CHO) and human 293 cells, insect cells such as Sf9 cells, yeast cells such as Saccharomyces cerevisiae, and bacteria such as E. coli. For each of these cell systems, a useful expression vector also includes an origin of replication to allow propagation in bacteria and a selectable marker such as the ß-lactamase antibiotic resistance gene to allow plasmid selection in bacteria. In addition, the vector may include a second selectable marker such as the neomycin phosphotransferase gene to allow selection in transfected eukaryotic host cells. Vectors for use in eukaryotic expression hosts require RNA processing elements such as 3' polyadenylation sequences if such are not part of the cDNA of interest.

Additionally, the vector contains promoters or enhancers which increase gene expression. Such promoters are host specific and include MMTV, SV40, and metallothionine promoters for CHO cells; trp, lac, tac and T7 promoters for bacterial hosts; and alpha factor, alcohol oxidase and PGH promoters for yeast. Transcription enhancers, such as the rous sarcoma virus enhancer, are used in mammalian host cells. Once homogeneous cultures of recombinant cells are obtained through standard culture methods, large quantities of recombinantly produced HEDG are recovered from the conditioned medium and analyzed using chromatographic methods known in the art.

VIII Isolation of Recombinant HEDG

HEDG is expressed as a chimeric protein with one or more additional polypeptide domains added to facilitate protein purification. Such purification facilitating domains include, but are not limited to, metal chelating peptides such as histidine-tryptophan modules that allow purification on immobilized metals, protein A domains that allow purification on immobilized immunoglobulin, and the domain utilized in the FLAGS extension/affinity purification system (Immunex Corp, Seattle WA). The inclusion of a cleavable linker sequence such as Factor XA or enterokinase (Invitrogen) between the purification domain and the HEDG sequence is useful to facilitate expression of HEDG.

IX Testing of Chimeric T7Gs

Functional chimeric T7Gs are constructed by combining the extracellular receptive sequences of a new isoform with the transmembrane and intracellular segments of a known isoform for test purposes. This concept was demonstrated by Kobilka et al (1988, Science 240:1310-1316) who created a series of chimeric α2-ß2 adrenergic receptors (AR) by inserting progressively greater amounts of α2-AR transmembrane sequence into ß2-AR. The

binding activity of known agonists changed as the molecule shifted from having more α2 than ß2 conformation, and intermediate constructs demonstrated mixed specificity. The specificity for binding antagonists, however, correlated with the source of the domain VII. The importance of T7G domain VII for ligand recognition was also found in chimeras utilizing two yeast α-factor receptors and is significant because the yeast receptors are classified as miscellaneous receptors. Thus, functional role of specific domains appears to be preserved throughout the T7G family regardless of category .

In parallel fashion, internal segments or cytoplasmic domains from a particular isoform are exchanged with the analogous domains of a known T7G and used to identify the structural determinants responsible for coupling the receptors to trimeric G-proteins (Dohlman et al (1991) Annu Rev Biochem 60:653-88). A chimeric receptor in which domains V, VI, and the intracellular connecting loop from ß2-AR were substituted into a2-AR was shown to bind ligands with a2-AR specificity, but to stimulate adenylate cyclase in the manner of ß2-AR. This demonstrates that for adrenergic-type receptors, G-protein recognition is present in domains V and VI and their connecting loop. The opposite situation was predicted and observed for a chimera in which the V->VI loop from α1-AR replaced the corresponding domain on ß2-AR and the resulting receptor bound ligands with ß2-AR specificity and activated G-protein-mediated phosphatidylinositol turnover in the α1-AR manner. Finally, chimeras constructed from muscarinic receptors also demonstrated that V->VI loop is the major determinant for specificity of G-protein activity (Bolander FF, supra).

Chimeric or modified T7Gs containing substitutions in the extracellular and transmembrane regions have shown that these portions of the receptor determine ligand binding specificity. For example, two Ser residues conserved in domain V of all adrenergic and D catecholamine T7G receptors are necessary for potent agonist activity. These serines are believed to form hydrogen bonds with the catechol moiety of the agonists within the T7G binding site. Similarly, an Asp residue present in domain III of all T7Gs which bind biogenic amines is believed to form an ion pair with the ligand amine group in the T7G binding site.

Functional, cloned T7Gs are expressed in heterologous expression systems and their biological activity assessed (eg Marullo et al (1988) Proc Natl Acad Sci 85:7551-55; King et al (1990) Science 250:121-23). One heterologous system introduces genes for a mammalian T7G and a mammalian G-protein into yeast cells. The T7G is shown to have appropriate ligand specificity and affinity and trigger appropriate biological activation--growth arrest and morphological changes--of the yeast cells.

An alternate procedure for testing chimeric receptors is based on the procedure utilizing the P_{2U} purinergic receptor (P_{2U}) as published by Erb et al (1993, Proc Natl Acad

Sci 90:10449-53). Function is asily tested in cultured K562 human leukemia cells because these cells lack P_{2U} r ceptors. K562 c lls are transf cted with expression vectors containing either normal or chimeric P_{2U} and loaded with fura-a, fluorescent probe for Ca^{++}. Activation of properly assembled and functional P_{2U} receptors with extracellular UTP or ATP mobilizes intracellular Ca^{++} which reacts with fura-a and is measured spectrofluorometrically.

As with the T7G receptors above, chimeric genes are created by combining sequences for extracellular receptive segments of any newly discovered T7G polypeptide with the nucleotides for the transmembrane and intracellular segments of the known P_{2U} molecule. Bathing the transfected K562 cells in microwells containing appropriate ligands triggers binding and fluorescent activity defining effectors of the T7G molecule. Once ligand and function are established, the P_{2U} system is useful for defining antagonists or inhibitors which block binding and prevent such fluorescent reactions.

X Production of HEDG Specific Antibodies

Two approaches are utilized to raise antibodies to HEDG, and each approach is useful for generating either polyclonal or monoclonal antibodies. In one approach, denatured protein from reverse phase HPLC separation is obtained in quantities up to 75 mg. This denatured protein is used to immunize mice or rabbits using standard protocols; about 100 micrograms are adequate for immunization of a mouse, while up to 1 mg might be used to immunize a rabbit. For identifying mouse hybridomas, the denatured protein is radioiodinated and used to screen potential murine B-cell hybridomas for those which produce antibody. This procedure requires only small quantities of protein, such that 20 mg is sufficient for labeling and screening of several thousand clones.

In the second approach, the amino acid sequence of an appropriate HEDG domain, as deduced from translation of the cDNA, is analyzed to determine regions of high antigenicity. Oligopeptides comprising appropriate hydrophilic regions, as illustrated in Figure 3, are synthesized and used in suitable immunization protocols to raise antibodies. Analysis to select appropriate epitopes is described by Ausubel FM et al (supra). The optimal amino acid sequences for immunization are usually at the C-terminus, the N-terminus and those intervening, hydrophilic regions of the polypeptide which are likely to be exposed to the external environment when the protein is in its natural conformation.

Typically, selected peptides, about 15 residues in length, are synthesized using an Applied Biosystems Peptide Synthesizer Model 431A using fmoc-chemistry and coupled to keyhole limpet hemocyanin (KLH; Sigma, St Louis MO) by reaction with M-maleimidobenzoyl-N- hydroxysuccinimid est r (MBS; Ausubel FM t al, supra). If necessary, a cysteine is

introduced at the N-terminus of the peptide to permit coupling to KLH. Rabbits are immunized with the peptide-KLH complex in compl te Freund's adjuvant. The resulting antisera are tested for antipeptide activity by binding the peptide to plastic, blocking with 1% bovine serum albumin, reacting with antisera, washing and reacting with labeled (radioactive or fluorescent), affinity purified, specific goat anti-rabbit IgG.

Hybridomas are prepared and screened using standard techniques. Hybridomas of interest are detected by screening with labeled HEDG to identify those fusions producing the monoclonal antibody with the desired specificity. In a typical protocol, wells of plates (FAST; Becton-Dickinson, Palo Alto CA) are coated during incubation with affinity purified, specific rabbit anti-mouse (or suitable antispecies Ig) antibodies at 10 mg/ml. The coated wells are blocked with 1% BSA, washed and incubated with supernatants from hybridomas. After washing the wells are incubated with labeled HEDG at 1 mg/ml. Supernatants with specific antibodies bind more labeled HEDG than is detectable in the background. Then clones producing specific antibodies are expanded and subjected to two cycles of cloning at limiting dilution. Cloned hybridomas are injected into pristane-treated mice to produce ascites, and monoclonal antibody is purified from mouse ascitic fluid by affinity chromatography on Protein A. Monoclonal antibodies with affinities of at least 10^8 M^{-1}, preferably 10^9 to 10^{10} or stronger, are typically made by standard procedures as described in Harlow and Lane (1988) Antibodies: A Laboratory Manual, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY; and in Goding (1986) Monoclonal Antibodies: Principles and Practice, Academic Press, New York City, both incorporated herein by reference.

XI Diagnostic Test Using HEDG Specific Antibodies

Particular HEDG antibodies are useful for investigating signal transduction and the diagnosis of infectious or hereditary conditions which are characterized by differences in the amount or distribution of HEDG or downstream products of an active signaling cascade. Since HEDG was found in a human rheumatoid library, it appears to be upregulated in cell types mainly involved in immune protection or defense.

Diagnostic tests for HEDG include methods utilizing antibody and a label to detect HEDG in human body fluids, membranes, cells, tissues or extracts of such. The polypeptides and antibodies of the present invention are used with or without modification. Frequently, the polypeptides and antibodies are labeled by joining them, either covalently or noncovalently, with a substance which provides for a detectable signal. A wide variety of labels and conjugation techniques are known and have been reported extensively in both the scientific and patent literature. Suitable labels include radionuclides, enzymes, substrates, cofactors, inhibitors, fluorescent agents, chemiluminescent agents, chromogenic agents, magnetic particles and the

like. Patents teaching the use of such labels include US Patent Nos. 3,817,837; 3,850,752; 3,939,350; 3,996,345; 4,277,437; 4,275,149; and 4,366,241. Also, recombinant immunoglobulins may be produced as shown in US Patent No. 4,816,567, incorporated herein by reference.

A variety of protocols for measuring soluble or membrane-bound HEDG, using either polyclonal or monoclonal antibodies specific for the protein, are known in the art. Examples include enzyme-linked immunosorbent assay (ELISA), radioimmunoassay (RIA) and fluorescent activated cell sorting (FACS). A two-site monoclonal-based immunoassay utilizing monoclonal antibodies reactive to two non-interfering epitopes on HEDG is preferred, but a competitive binding assay may be employed. These assays are described, among other places, in Maddox, DE et al (1983, J Exp Med 158:1211f).

XII Purification of Native HEDG Using Specific Antibodies

Native or recombinant HEDG is purified by immunoaffinity chromatography using antibodies specific for HEDG. In general, an immunoaffinity column is constructed by covalently coupling the anti-TRH antibody to an activated chromatographic resin.

Polyclonal immunoglobulins are prepared from immune sera either by precipitation with ammonium sulfate or by purification on immobilized Protein A (Pharmacia LKB Biotechnology, Piscataway NJ). Likewise, monoclonal antibodies are prepared from mouse ascites fluid by ammonium sulfate precipitation or chromatography on immobilized Protein A. Partially purified immunoglobulin is covalently attached to a chromatographic resin such as CnBr-activated Sepharose (Pharmacia LKB Biotechnology). The antibody is coupled to the resin, the resin is blocked, and the derivative resin is washed according to the manufacturer's instructions.

Such immunoaffinity columns are utilized in the purification of HEDG by preparing a fraction from cells containing HEDG in a soluble form. This preparation is derived by solubilization of whole cells or of a subcellular fraction obtained via differential centrifugation (with or without addition of detergent) or by other methods well known in the art. Alternatively, soluble HEDG containing a signal sequence is secreted in useful quantity into the medium in which the cells are grown.

A soluble HEDG-containing preparation is passed over the immunoaffinity column, and the column is washed under conditions that allow the preferential absorbance of HEDG (eg, high ionic strength buffers in the presence of detergent). Then, the column is eluted under conditions that disrupt antibody/protein binding (eg, a buffer of pH 2-3 or a high concentration of a chaotrope such as urea or thiocyanate ion), and HEDG is collected.

XIII Drug Screening

This invention is particularly useful for screening therapeutic compounds by using HEDG or binding fragments thereof in any of a variety of drug scre ning techniques. The polypeptide or fragment employed in such a test is either free in solution, affixed to a solid support, borne on a cell surface or located intracellularly. One method of drug screening utilizes eukaryotic or prokaryotic host cells which are stably transformed with recombinant nucleic acids expressing the polypeptide or fragment. Drugs are screened against such transformed cells in competitive binding assays. Such cells, either in viable or fixed form, are used for standard binding assays. One measures, for example, the formation of complexes between HEDG and the agent being tested. Alternatively, one examines the diminution in complex formation between HEDG and a receptor caused by the agent being tested.

Thus, the present invention provides methods of screening for drugs or any other agents which affect signal transduction. These methods, well known in the art, comprise contacting such an agent with HEDG polypeptide or a fragment thereof and assaying (i) for the presence of a complex between the agent and the HEDG polypeptide or fragment, or (ii) for the presence of a complex between the HEDG polypeptide or fragment and the cell. In such competitive binding assays, the HEDG polypeptide or fragment is typically labeled. After suitable incubation, free HEDG polypeptide or fragment is separated from that present in bound form, and the amount of free or uncomplexed label is a measure of the ability of the particular agent to bind to HEDG or to interfere with the HEDG and agent complex.

Another technique for drug screening provides high throughput screening for compounds having suitable binding affinity to the HEDG polypeptides and is described in detail in European Patent Application 84/03564, published on September 13, 1984, incorporated herein by reference. Briefly stated, large numbers of different small peptide test compounds are synthesized on a solid substrate, such as plastic pins or some other surface. The peptide test compounds are reacted with HEDG polypeptide and washed. Bound HEDG polypeptide is then detected by methods well known in the art. Purified HEDG are also coated directly onto plates for use in the aforementioned drug screening techniques. In addition, non-neutralizing antibodies are used to capture the peptide and immobilize it on the solid support.

This invention also contemplates the use of competitive drug screening assays in which neutralizing antibodies capable of binding HEDG specifically compete with a test compound for binding to HEDG polypeptides or fragments thereof. In this manner, the antibodies are used to detect the presence of any peptide which shares one or more antigenic determinants with HEDG.

XIV Rational Drug Design

The goal of rational drug design is to produce structural analogs of biologically active polypeptides of interest or of small molecules with which they interact, agonists, antagonists,

or inhibitors. Any of these examples are used to fashion drugs which are more active or stable forms of the polypeptide or which enhance or interfere with the function of a polypeptide in vivo (eg Hodgson J (1991) Bio/Technology 9:19-21, incorporated herein by reference).

In one approach, the three-dimensional structure of a protein of interest, or of a protein-inhibitor complex, is determined by x-ray crystallography, by computer modeling or, most typically, by a combination of the two approaches. Both the shape and charges of the polypeptide must be ascertained to elucidate the structure and to determine active site(s) of the molecule. Less often, useful information regarding the structure of a polypeptide is gained by modeling based on the structure of homologous proteins. In both cases, relevant structural information is used to design efficient inhibitors. Useful examples of rational drug design includes molecules which have improved activity or stability as shown by Braxton S and Wells JA (1992, Biochemistry 31:7796- 7801) or which act as inhibitors, agonists, or antagonists of native peptides as shown by Athauda SB et al (1993 J Biochem 113:742-46), incorporated herein by reference.

It is also possible to isolate a target-specific antibody, selected by functional assay, as described above, and then to solve its crystal structure. This approach, in principle, yields a pharmacore upon which subsequent drug design is based. It is possible to bypass protein crystallography altogether by generating anti-idiotypic antibodies (anti-ids) to a functional, pharmacologically active antibody. As a mirror image of a mirror image, the binding site of the anti-ids is expected to be an analog of the original receptor. The anti-id is then used to identify and isolate peptides from banks of chemically or biologically produced peptides. The isolated peptides then act as the pharmacore.

By virtue of the present invention, sufficient amount of polypeptide are made available to perform such analytical studies as X-ray crystallography. In addition, knowledge of the HEDG amino acid sequence provided herein provides guidance to those employing computer modeling techniques in place of or in addition to x-ray crystallography.

XV Identification of Other Members of the Signal Transduction Complex

The inventive purified HEDG is a research tool for identification, characterization and purification of interacting G or other signal transduction pathway proteins. Radioactive labels are incorporated into a selected HEDG domain by various methods known in the art and used in vitro to capture interacting molecules. A preferred method involves labeling the primary amino groups in HEDG with ^{125}I Bolton-Hunter reagent (Bolton, AE and Hunter, WM (1973) Biochem J 133: 529). This reagent has been used to label various molecules without concomitant loss of biological activity (Hebert CA et al (1991) J Biol Chem 266: 18989; McColl S et al (1993) J Immunol 150:4550-4555).

Labeled HEDG is useful as a reagent for the purification of molecules with which it interacts. In on embodiment of affinity purification, membrane-bound HEDG is covalently coupled to a chromatography column. Cell-free extract derived from synovial cells or putative target cells is passed over the column, and molecules with appropriate affinity bind to HEDG. HEDG-complex is recovered from the column, and the HEDG-binding ligand disassociated and subjected to N-terminal protein sequencing. This aa sequence is then used to identify the captured molecule or to design degenerate oligonucleotide probes for cloning the relevant gene from an appropriate cDNA library.

In an alternate method, antibodies are raised against HEDG, specifically monoclonal antibodies. The monoclonal antibodies are screened to identify those which inhibit the binding of labeled HEDG. These monoclonal antibodies are then used therapeutically.

XVI Use and Administration of Antibodies, Inhibitors, or Antagonists

Antibodies, inhibitors, or antagonists of HEDG (or other treatments to limit signal transduction, LST) provide different effects when administered therapeutically. LSTs are formulated in a nontoxic, inert, pharmaceutically acceptable aqueous carrier medium preferably at a pH of about 5 to 8, more preferably 6 to 8, although pH may vary according to the characteristics of the antibody, inhibitor, or antagonist being formulated and the condition to be treated. Characteristics of LSTs include solubility of the molecule, half-life and antigenicity/immunogenicity. These and other characteristics aid in defining an effective carrier. Native human proteins are preferred as LSTs, but organic or synthetic molecules resulting from drug screens are equally effective in particular situations.

LSTs are delivered by known routes of administration including but not limited to topical creams and gels; transmucosal spray and aerosol; transdermal patch and bandage; injectable, intravenous and lavage formulations; and orally administered liquids and pills particularly formulated to resist stomach acid and enzymes. The particular formulation, exact dosage, and route of administration is determined by the attending physician and varies according to each specific situation.

Such determinations are made by considering multiple variables such as the condition to be treated, the LST to be administered, and the pharmacokinetic profile of a particular LST. Additional factors which are taken into account include severity of the disease state, patient's age, weight, gender and diet, time and frequency of LST administration, possible combination with other drugs, reaction sensitivities, and tolerance/response to therapy. Long acting LST formulations might be administered every 3 to 4 days, every week, or once every two weeks depending on half-life and clearance rate of the particular LST.

Normal dosage amounts vary from 0.1 to 100,000 micrograms, up to a total dose of

about 1 g, depending upon the rout of administration. Guidance as to particular dosages and methods of delivery is provided in the literature; see US Patent Nos. 4,657,760; 5,206,344; or 5,225,212. Those skilled in the art employ different formulations for different LSTs. Administration to cells such as nerve cells necessitates delivery in a manner different from that to other cells such as vascular endothelial cells.

It is contemplated that abnormal signal transduction, trauma, or diseases which trigger HEDG activity are treatable with LSTs. These conditions or diseases are specifically diagnosed by the tests discussed above, and such testing should be performed in suspected cases of viral, bacterial or fungal infections; allergic responses; mechanical injury associated with trauma; hereditary diseases; lymphoma or carcinoma; or other conditions which activate the genes of lymphoid tissues.

All publications and patents mentioned in the above specification are herein incorporated by reference. Various modifications and variations of the described method and system of the invention will be apparent to those skilled in the art without departing from the scope and spirit of the invention. Although the invention has been described in connection with specific preferred embodiments, it should be understood that the invention as claimed should not be unduly limited to such specific embodiments. Indeed, various modifications of the above-described modes for carrying out the invention which are obvious to those skilled in the field of molecular biology or related fields are intended to be within the scope of the following claims.

SEQUENCE LISTING

(1) GENERAL INFORMATION

(i) APPLICANT: Coleman, Roger
Guegler, Karl J.
Au-Young, Janice
Bandman, Olga
Seilhamer, Jeffrey J.

(ii) TITLE OF THE INVENTION: A NOVEL HUMAN EDG-2 RECEPTOR HOMOLOG

(iii) NUMBER OF SEQUENCES: 6

(iv) CORRESPONDENCE ADDRESS:
(A) ADDRESSEE: Incyte Pharmaceuticals, Inc.
(B) STREET: 3174 Porter Drive
(C) CITY: Palo Alto
(D) STATE: CA
(E) COUNTRY: USA
(F) ZIP: 94304

(v) COMPUTER READABLE FORM:
(A) MEDIUM TYPE: Diskette
(B) COMPUTER: IBM Compatible
(C) OPERATING SYSTEM: DOS
(D) SOFTWARE: FastSEQ Version 1.5

(vi) CURRENT APPLICATION DATA:
(A) APPLICATION NUMBER: TO BE ASSIGNED
(B) FILING DATE: 20-JUN-1996
(C) CLASSIFICATION:

(vii) PRIOR APPLICATION DATA:
(A) APPLICATION NUMBER: 60/000,352
(B) FILING DATE: 20-JUN-1995

(vii) PRIOR APPLICATION DATA:
(A) APPLICATION NUMBER: 08/567,817
(B) FILING DATE: 06-DEC-1995

(viii) ATTORNEY/AGENT INFORMATION:
(A) NAME: Glaister, Debra J.
(B) REGISTRATION NUMBER: 33,888
(C) REFERENCE/DOCKET NUMBER: PF-0042 PCT

(ix) TELECOMMUNICATION INFORMATION:
(A) TELEPHONE: 415-855-0555
(B) TELEFAX: 415-845-4166
(C) TELEX:

(2) INFORMATION FOR SEQ ID NO:1:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1875 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(vii) IMMEDIATE SOURCE:
(A) LIBRARY: Rheumatoid Synovium
(B) CLONE: 80853

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

```
GGCAGGTACG GCCGGATTCC CGGGTCGACC ACGCGTCCGC TCTCAAGGGA ACAGCTCCTG      60
CCCAGGTCTG TGGGTACTCA GCATGGATAT CAGTCTCCCT GTGAGTGATG GGAAAGAACT     120
AGCAGAGGTG GACGTCTGAT TTATGAAGCT CCCCATCCAC CTATCTGAGT ACCTGACTTC     180
TCAGGACTGA CACCTACAGC ATCAGGTACA CAGCTTCTCC TAGCATGACT TCGATCTGAT     240
CACCACACAA GAAAATTTGT CTCCCGTAGT TCTGGGGCGT GTTCACCACC TACAACCACA     300
GAGCTGTCAT GGCTGCCATC TCTACTTCCA TCCCTGTAAT TTCACAGCCC CAGTTCACAG     360
CCATGAATGA ACCACAGTGC TTCTACAACG AGTCCATTGC CTTCTTTTAT AACCGAAGTG     420
GAAAGCATCT TGCCACAGAA TGGAACACAG TCAGCAAGCT GGTGATGGGA CTTGGAATCA     480
CTGTTTGTAT CTTCATCATG TTGGCCAACC TATTGGTCAT GGTGGCAATC TATGTCAACC     540
GCCGCTTCCA TTTTCCTATT TATTACCTAA TGGCTAATCT GGCTGCTGCA GACTTCTTTG     600
CTGGGTTGGC CTACTTCTAT CTCATGTTCA ACACAGGACC CAATACTCGG AGACTGACTG     660
TTAGCACATG GCTCCTTCGT CAGGGCCTCA TTGACACCAG CCTGACGGCA TCTGTGGCCA     720
ACTTACTGGC TATTGCAATC GAGAGGCACA TTACGGTTTT CCGCATGCAG CTCCACACAC     780
GGATGAGCAA CCGGCGGGTA GTGGTGGTCA TTGTGGTCAT CTGGACTATG GCCATCGTTA     840
TGGGTGCTAT ACCCAGTGTG GGCTGGAACT GTATCTGTGA TATTGAAAAT TGTTCCAACA     900
TGGCACCCCT CTACAGTGAC TCTTACTTAG TCTTCTGGGC CATTTTCAAC TTGGTGACCT     960
TTGTGGTAAT GGTGGTTCTC TATGCTCACA TCTTTGGCTA TGTTCGCCAG AGGACTATGA    1020
GAATGTCTCG GCATAGTTCT GGACCCCGGC GGAATCGGGA TACCATGATG AGTCTTCTGA    1080
AGACTGTGGT CATTGTGCTT GGGGGCTTTA TCATCTGCTG GACTCCTGGA TTGGTTTTGT    1140
TACTTCTAGA CGTGTGCTGT CCACAGTGCG ACGTGCTGGC CTATGAGAAA TTCTTCCTTC    1200
TCCTTGCTGA ATTCAACTCT GCCATGAACC CCATCATTTA CTCCTACCGT GACAAAGAAA    1260
TGAGCGCCAC CTTTAGACAG ATCCTCTGCT GCCAGCGCAG TGAGAACCCC ACCGCCCCCA    1320
CAGAAGGCTC AGACCGCTCG GCTTCCTCCC TCAACCACAC CATCTTGGCT GGAGTTCACA    1380
GCAATGACCA CTCTGTGGTT TAGAACGGAA ACTGAGATGA GGAACCAGCC GTCCTCTCTT    1440
GGAGGATAAA CAAGCCTCCC CCTACCCAAT TGCCAGGGCA AGGTGGGGTG TGAGAGAGGA    1500
GAAAAGTCAA CTCATGTACT TAAACACTAA CCAATGACAG TATTTGTTCC TGGACCCCAC    1560
AAGACTTGAT ATATATTGAA AATTAGCTTA TGTGACAACC CTCATCTTGA TCCCCATCCC    1620
TTCTGAAAGT AGGAAGTTGG AGCTCTTGCA ATGGAATTCA AGAACAGACT CTGGAGTGTC    1680
CATTTAGACT ACACTAACTA GACTTTTAAA AGATTTTGTG TGGTTTGGTG CAAGTCAGAA    1740
TAAATTCTGG CTAGTTGAAT CCACAACTTC ATTTATATAC AGGCTTCCCT TTTTTATTTT    1800
TAAAGGATAC GTTTCACTTA ATAAACACGT TTATGCCTAA AAAAAAAAAA AAAAAAAAAA    1860
AAAAAAAAAA AAAAC                                                     1875
```

(2) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 364 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(vii) IMMEDIATE SOURCE:
(A) LIBRARY: Rheumatoid Synovium
(B) CLONE: 80853

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

```
Met Ala Ala Ile Ser Thr Ser Ile Pro Val Ile Ser Gln Pro Gln Phe
 1               5                  10                  15
```

```
Thr Ala Met Asn Glu Pro Gln Cys Phe Tyr Asn Glu Ser Ile Ala Phe
            20                  25                  30
Phe Tyr Asn Arg Ser Gly Lys His Leu Ala Thr Glu Trp Asn Thr Val
        35                  40                  45
Ser Lys Leu Val Met Gly Leu Gly Ile Thr Val Cys Ile Phe Ile Met
    50                  55                  60
Leu Ala Asn Leu Leu Val Met Val Ala Ile Tyr Val Asn Arg Arg Phe
65                  70                  75                  80
His Phe Pro Ile Tyr Tyr Leu Met Ala Asn Leu Ala Ala Ala Asp Phe
                85                  90                  95
Phe Ala Gly Leu Ala Tyr Phe Tyr Leu Met Phe Asn Thr Gly Pro Asn
            100                 105                 110
Thr Arg Arg Leu Thr Val Ser Thr Trp Leu Leu Arg Gln Gly Leu Ile
        115                 120                 125
Asp Thr Ser Leu Thr Ala Ser Val Ala Asn Leu Leu Ala Ile Ala Ile
    130                 135                 140
Glu Arg His Ile Thr Val Phe Arg Met Gln Leu His Thr Arg Met Ser
145                 150                 155                 160
Asn Arg Arg Val Val Val Val Ile Val Val Ile Trp Thr Met Ala Ile
                165                 170                 175
Val Met Gly Ala Ile Pro Ser Val Gly Trp Asn Cys Ile Cys Asp Ile
            180                 185                 190
Glu Asn Cys Ser Asn Met Ala Pro Leu Tyr Ser Asp Ser Tyr Leu Val
        195                 200                 205
Phe Trp Ala Ile Phe Asn Leu Val Thr Phe Val Val Met Val Val Leu
    210                 215                 220
Tyr Ala His Ile Phe Gly Tyr Val Arg Gln Arg Thr Met Arg Met Ser
225                 230                 235                 240
Arg His Ser Ser Gly Pro Arg Arg Asn Arg Asp Thr Met Met Ser Leu
                245                 250                 255
Leu Lys Thr Val Val Ile Val Leu Gly Gly Phe Ile Ile Cys Trp Thr
            260                 265                 270
Pro Gly Leu Val Leu Leu Leu Leu Asp Val Cys Cys Pro Gln Cys Asp
        275                 280                 285
Val Leu Ala Tyr Glu Lys Phe Phe Leu Leu Leu Ala Glu Phe Asn Ser
    290                 295                 300
Ala Met Asn Pro Ile Ile Tyr Ser Tyr Arg Asp Lys Glu Met Ser Ala
305                 310                 315                 320
Thr Phe Arg Gln Ile Leu Cys Cys Gln Arg Ser Glu Asn Pro Thr Ala
                325                 330                 335
Pro Thr Glu Gly Ser Asp Arg Ser Ala Ser Ser Leu Asn His Thr Ile
            340                 345                 350
Leu Ala Gly Val His Ser Asn Asp His Ser Val Val
        355                 360
```

(2) INFORMATION FOR SEQ ID NO:3:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 393 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(vii) IMMEDIATE SOURCE:
(A) LIBRARY: GenBank
(B) CLONE:U18405

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

```
Met Ala Ala Ala Ser Thr Ser Ser Pro Val Val Ser Gln Pro Gln Phe
 1               5                  10                  15
Thr Ala Met Asn Glu Pro Gln Cys Phe Tyr Asn Glu Ser Ile Ala Phe
            20                  25                  30
Phe Tyr Asn Arg Ser Gly Lys Tyr Leu Ala Thr Glu Trp Asn Thr Val
        35                  40                  45
Ser Lys Leu Val Met Gly Leu Gly Ile Thr Val Cys Ile Phe Ile Met
    50                  55                  60
Leu Ala Asn Leu Leu Val Met Val Ala Ile Tyr Val Asn Arg Arg Phe
65                  70                  75                  80
His Phe Pro Ile Tyr Tyr Leu Met Ala Asn Leu Ala Ala Ala Asp Phe
                85                  90                  95
Phe Ala Gly Leu Ala Tyr Phe Tyr Leu Met Phe Asn Thr Gly Pro Asn
            100                 105                 110
Thr Arg Arg Leu Thr Val Ser Thr Trp Leu Leu Arg Gln Gly Leu Ile
        115                 120                 125
Asp Thr Thr Val Thr Ala Ser Val Ala Asn Leu Leu Ala Ile Ala Ile
    130                 135                 140
Glu Arg His Ile Thr Val Phe Arg Met Gln Leu His Thr Arg Met Ser
145                 150                 155                 160
Asn Arg Arg Val Val Val Val Ile Val Val Ile Trp Thr Met Ala Ile
                165                 170                 175
Val Met Gly Ala Ile Pro Ser Val Gly Trp Asn Cys Ile Cys Asp Ile
            180                 185                 190
Glu Asn Cys Ser Asn Met Ala Pro Leu Tyr Ser Asp Ser Tyr Leu Val
        195                 200                 205
Phe Trp Ala Ile Phe Asn Leu Val Thr Phe Val Val Met Val Val Leu
    210                 215                 220
Tyr Ala His Ile Phe Gly Tyr Val Arg Gln Arg Thr Met Arg Met Ser
225                 230                 235                 240
Arg His Ser Ser Gly Pro Arg Arg Asn Arg Asp Thr Met Met Ser Leu
                245                 250                 255
Leu Lys Thr Val Val Ile Val Leu Gly Ala Phe Ile Ile Cys Trp Thr
            260                 265                 270
Pro Gly Leu Val Leu Leu Leu Leu Asp Val Cys Cys Pro Gln Cys Asp
        275                 280                 285
Val Leu Ala Tyr Glu Lys Phe Phe Leu Leu Leu Ala Glu Phe Asn Ser
    290                 295                 300
Ala Met Asn Pro Ile Ile Tyr Ser Tyr Arg Asp Lys Glu Met Ser Ala
305                 310                 315                 320
Thr Phe Arg Gln Ile Leu Cys Cys Gln Arg Ser Glu Asn Thr Ser Gly
                325                 330                 335
Pro Thr Glu Gly Ser Asp Arg Ser Ala Ser Ser Leu Asn His Thr Ile
            340                 345                 350
Leu Ala Gly Val His Ser Asn Asp His Ser Val Phe Arg Lys Glu Thr
        355                 360                 365
Lys Met Arg Gly Gly His His Leu Leu Arg Asp Glu Gln Pro Pro Pro
    370                 375                 380
Pro Glu Arg Pro Gly Gln Gly Arg Val
385                 390
```

(2) INFORMATION FOR SEQ ID NO:4:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 381 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(vii) IMMEDIATE SOURCE:
(A) LIBRARY: GenBank
(B) CLONE: 119130

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

```
Met Gly Pro Thr Ser Val Pro Leu Val Lys Ala His Arg Ser Ser Val
 1               5                  10                  15
Ser Asp Tyr Val Asn Tyr Asp Ile Ile Val Arg His Tyr Asn Tyr Thr
            20                  25                  30
Gly Lys Leu Asn Ile Ser Ala Asp Lys Glu Asn Ser Ile Lys Leu Thr
        35                  40                  45
Ser Val Val Phe Ile Leu Ile Cys Cys Phe Ile Ile Leu Glu Asn Ile
    50                  55                  60
Phe Val Leu Leu Thr Ile Trp Lys Thr Lys Lys Phe His Arg Pro Met
65                  70                  75                  80
Tyr Tyr Phe Ile Gly Asn Leu Ala Leu Ser Asp Leu Leu Ala Gly Val
                85                  90                  95
Ala Tyr Thr Ala Asn Leu Leu Leu Ser Gly Ala Thr Thr Tyr Lys Leu
            100                 105                 110
Thr Pro Ala Gln Trp Phe Leu Arg Glu Gly Ser Met Phe Val Ala Leu
        115                 120                 125
Ser Ala Ser Val Phe Ser Leu Leu Ala Ile Ala Ile Glu Arg Tyr Ile
    130                 135                 140
Thr Met Leu Lys Met Lys Leu His Asn Gly Ser Asn Asn Phe Arg Leu
145                 150                 155                 160
Phe Leu Leu Ile Ser Ala Cys Trp Val Ile Ser Leu Ile Leu Gly Gly
                165                 170                 175
Leu Pro Ile Met Gly Trp Asn Cys Ile Ser Ala Leu Ser Ser Cys Ser
            180                 185                 190
Thr Val Leu Pro Leu Tyr His Lys His Tyr Ile Leu Phe Cys Thr Thr
        195                 200                 205
Val Phe Thr Leu Leu Leu Leu Ser Ile Val Ile Leu Tyr Cys Arg Ile
    210                 215                 220
Tyr Ser Leu Val Arg Thr Arg Ser Arg Arg Leu Thr Phe Arg Lys Asn
225                 230                 235                 240
Ile Ser Lys Ala Ser Arg Ser Ser Glu Asn Val Ala Leu Leu Lys Thr
                245                 250                 255
Val Ile Ile Val Leu Ser Val Phe Ile Ala Cys Trp Ala Pro Leu Phe
            260                 265                 270
Ile Leu Leu Leu Leu Asp Val Gly Cys Lys Val Lys Thr Cys Asp Ile
        275                 280                 285
Leu Phe Arg Ala Glu Tyr Phe Leu Val Leu Ala Val Leu Asn Ser Gly
    290                 295                 300
Thr Asn Pro Ile Ile Tyr Thr Leu Thr Asn Lys Glu Met Arg Arg Ala
305                 310                 315                 320
Phe Ile Arg Ile Met Ser Cys Cys Lys Cys Pro Ser Gly Asp Ser Ala
                325                 330                 335
Gly Lys Phe Lys Arg Pro Ile Ile Ala Gly Met Glu Phe Ser Arg Ser
            340                 345                 350
Lys Ser Asp Asn Ser Ser His Pro Gln Lys Asp Glu Gly Asp Asn Pro
        355                 360                 365
Glu Thr Ile Met Ser Ser Gly Asn Val Asn Ser Ser Ser
    370                 375                 380
```

(2) INFORMATION FOR SEQ ID NO:5:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 24 base pairs
(B) TYPE: nucleic acid

(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(vii) IMMEDIATE SOURCE:
(A) Oligomer R

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

TCATCTTGAT CCCCATCCCT TCTG 24

(2) INFORMATION FOR SEQ ID NO:6:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 24 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(vii) IMMEDIATE SOURCE:
(A) Oligomer F

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

AGTCTCCGAG TATTGGGTCC TGTG 24

CLAIMS

1. A purified polynucleotide comprising a nucleic acid sequence encoding the polypeptide of SEQ ID NO:2, or the complement of said polynucleotide.

2. The polynucleotide of Claim 1 comprising the nucleic acid sequence for (hedg) of SEQ ID NO:1.

3. An antisense molecule comprising the complement of the polynucleotide of Claim 2 or a portion thereof.

4. A pharmaceutical composition comprising the antisense molecule of Claim 3 and a pharmaceutically acceptable excipient.

5. A method of treating a subject with a condition associated with altered hedg expression comprising administering an effective amount of the pharmaceutical composition of Claim 4 to the subject.

6. A diagnostic composition comprising an oligomer of the polynucleotide of Claim 2.

7. A diagnostic test for a condition associated with altered hedg expression comprising the steps of:

a) providing a biological sample;

b) combining the biological sample and the diagnostic composition of Claim 6;

c) allowing hybridization to occur between the biological sample and the diagnostic composition under suitable conditions;

d) measuring the amount of hybridization to obtain a sample value; and

e) comparing the sample value with standard values to determine whether hedg expression is altered.

8. An expression vector comprising the polynucleotide of Claim 1.

9. A host cell transformed with the expression vector of Claim 8.

10. A method for producing a polypeptide, said method comprising the steps of:

a) culturing the host cell of Claim 9 under conditions suitable for the expression of the polypeptide; and

b) recovering the polypeptide from the host cell culture.

11. A purified polypeptide (HEDG) comprising the amino acid sequence of SEQ ID NO:2.

12. A diagnostic composition comprising the polypeptide of Claim 11 or a portion thereof.

13. A pharmaceutical composition comprising the polypeptide of Claim 11 and a pharmaceutically acceptable excipient.

14. A method of treating a subject with a condition associated with altered HEDG expression comprising administering an effective amount of the pharmaceutical composition of Claim 13 to the subject.

15. An antibody specific for the purified polypeptide of Claim 11, or portion thereof.

16. A diagnostic composition comprising the antibody of Claim 15.

17. A diagnostic test for a condition associated with altered HEDG expression comprising the steps of:

a) providing a biological sample;

b) combining the biological sample and the antibody of Claim 15 under conditions suitable for complex formation;

c) measuring the amount of complex formation between HEDG and the antibody to obtain a sample amount; and

d) comparing the amount of complex formation in the sample with standard amounts of complex formation, wherein a variation between the sample amount and standard amounts of complex formation establishes the presence of the condition.

18. A method of screening a plurality of compounds for specific binding affinity with the polypeptide of Claim 11 or any portion thereof comprising the steps of:

a) providing a plurality of compounds;

b) combining HEDG with each of a plurality of compounds for a time sufficient to allow binding under suitable conditions; and

c) detecting binding of HEDG to each of the plurality of compounds, thereby identifying the compounds which specifically bind HEDG.

19. A pharmaceutical composition comprising a compound of Claim 18 and a pharmaceutically acceptable excipient.

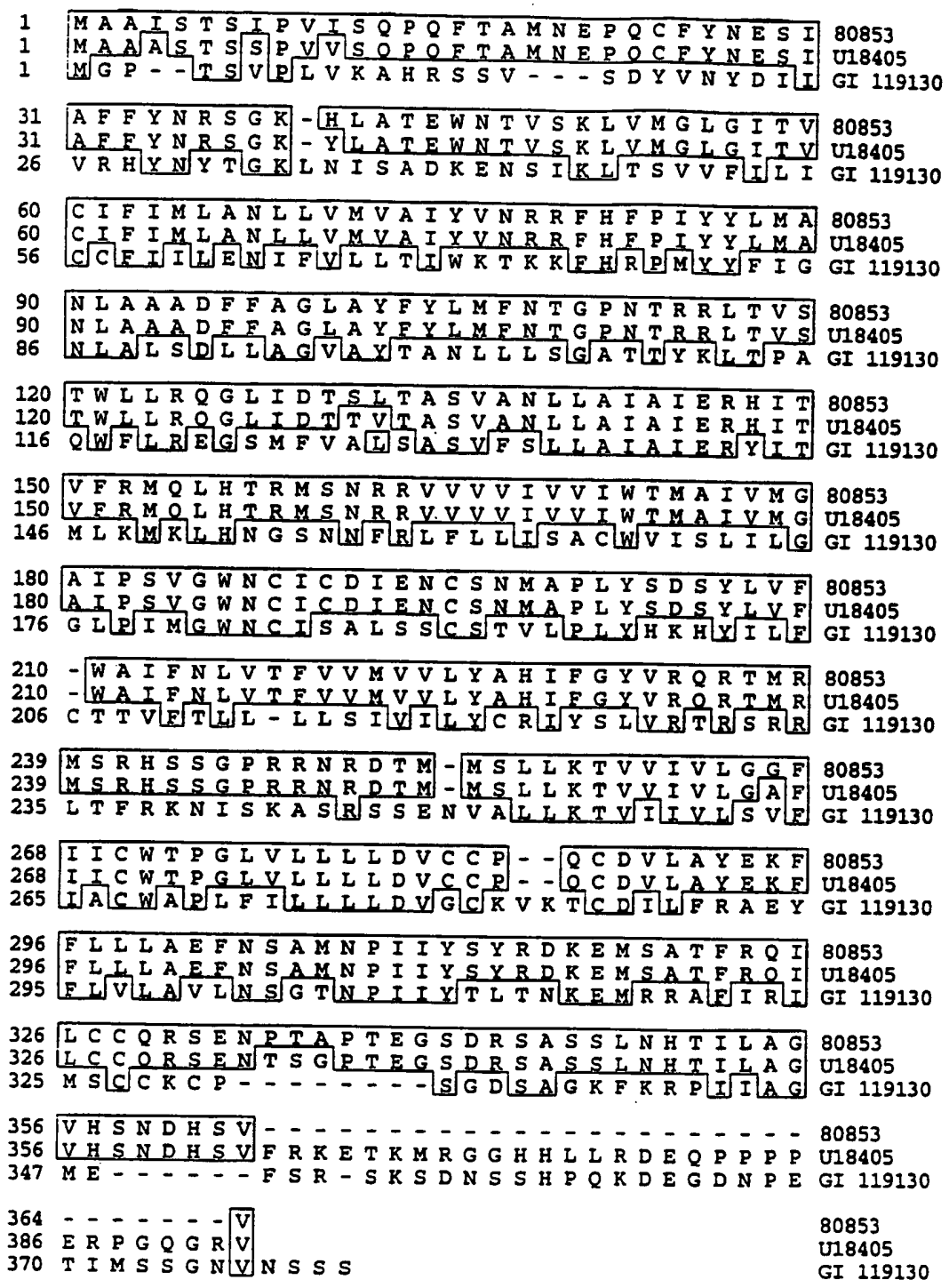
20. A method of treating a subject with a condition associated with altered HEDG expression comprising administering an effective amount of the pharmaceutical composition of Claim 19 to the subject.

```
                9         18         27         36         45         54
5' ATG GCT GCC ATC TCT ACT TCC ATC CCT GTA ATT TCA CAG CCC CAG TTC ACA GCC
   M   A   A   I   S   T   S   I   P   V   I   S   Q   P   Q   F   T   A

               63         72         81         90         99        108
   ATG AAT GAA CCA CAG TGC TTC TAC AAC GAG TCC ATT GCC TTC TTT TAT AAC CGA
   M   N   E   P   Q   C   F   Y   N   E   S   I   A   F   F   Y   N   R

              117        126        135        144        153        162
   AGT GGA AAG CAT CTT GCC ACA GAA TGG AAC ACA GTC AGC AAG CTG GTG ATG GGA
   S   G   K   H   L   A   T   E   W   N   T   V   S   K   L   V   M   G

              171        180        189        198        207        216
   CTT GGA ATC ACT GTT TGT ATC TTC ATC ATG TTG GCC AAC CTA TTG GTC ATG GTG
   L   G   I   T   V   C   I   F   I   M   L   A   N   L   L   V   M   V

              225        234        243        252        261        270
   GCA ATC TAT GTC AAC CGC CGC TTC CAT TTT CCT ATT TAT TAC CTA ATG GCT AAT
   A   I   Y   V   N   R   R   F   H   F   P   I   Y   Y   L   M   A   N

              279        288        297        306        315        324
   CTG GCT GCT GCA GAC TTC TTT GCT GGG TTG GCC TAC TTC TAT CTC ATG TTC AAC
   L   A   A   A   D   F   F   A   G   L   A   Y   F   Y   L   M   F   N

              333        342        351        360        369        378
   ACA GGA CCC AAT ACT CGG AGA CTG ACT GTT AGC ACA TGG CTC CTT CGT CAG GGC
   T   G   P   N   T   R   R   L   T   V   S   T   W   L   L   R   Q   G

              387        396        405        414        423        432
   CTC ATT GAC ACC AGC CTG ACG GCA TCT GTG GCC AAC TTA CTG GCT ATT GCA ATC
   L   I   D   T   S   L   T   A   S   V   A   N   L   L   A   I   A   I

              441        450        459        468        477        486
   GAG AGG CAC ATT ACG GTT TTC CGC ATG CAG CTC CAC ACA CGG ATG AGC AAC CGG
   E   R   H   I   T   V   F   R   M   Q   L   H   T   R   M   S   N   R

              495        504        513        522        531        540
   CGG GTA GTG GTG GTC ATT GTG GTC ATC TGG ACT ATG GCC ATC GTT ATG GGT GCT
   R   V   V   V   V   I   V   V   I   W   T   M   A   I   V   M   G   A

              549        558        567        576        585        594
   ATA CCC AGT GTG GGC TGG AAC TGT ATC TGT GAT ATT GAA AAT TGT TCC AAC ATG
   I   P   S   V   G   W   N   C   I   C   D   I   E   N   C   S   N   M

              603        612        621        630        639        648
   GCA CCC CTC TAC AGT GAC TCT TAC TTA GTC TTC TGG GCC ATT TTC AAC TTG GTG
   A   P   L   Y   S   D   S   Y   L   V   F   W   A   I   F   N   L   V

              657        666        675        684        693        702
   ACC TTT GTG GTA ATG GTG GTT CTC TAT GCT CAC ATC TTT GGC TAT GTT CGC CAG
   T   F   V   V   M   V   V   L   Y   A   H   I   F   G   Y   V   R   Q

              711        720        729        738        747        756
   AGG ACT ATG AGA ATG TCT CGG CAT AGT TCT GGA CCC CGG CGG AAT CGG GAT ACC
   R   T   M   R   M   S   R   H   S   S   G   P   R   R   N   R   D   T

              765        774        783        792        801        810
   ATG ATG AGT CTT CTG AAG ACT GTG GTC ATT GTG CTT GGG GGC TTT ATC ATC TGC
   M   M   S   L   L   K   T   V   V   I   V   L   G   G   F   I   I   C

              819        828        837        846        855        864
   TGG ACT CCT GGA TTG GTT TTG TTA CTT CTA GAC GTG TGC TGT CCA CAG TGC GAC
   W   T   P   G   L   V   L   L   L   L   D   V   C   C   P   Q   C   D
```

FIGURE 1A

```
                873                 882                 891                 900                 909                 918
GTG CTG GCC TAT GAG AAA TTC TTC CTT CTC CTT GCT GAA TTC AAC TCT GCC ATG
V   L   A   Y   E   K   F   F   L   L   L   A   E   F   N   S   A   M

                927                 936                 945                 954                 963                 972
AAC CCC ATC ATT TAC TCC TAC CGT GAC AAA GAA ATG AGC GCC ACC TTT AGA CAG
N   P   I   I   Y   S   Y   R   D   K   E   M   S   A   T   F   R   Q

                981                 990                 999                1008                1017                1026
ATC CTC TGC TGC CAG CGC AGT GAG AAC CCC ACC GCC CCC ACA GAA GGC TCA GAC
I   L   C   C   Q   R   S   E   N   P   T   A   P   T   E   G   S   D

               1035                1044                1053                1062                1071                1080
CGC TCG GCT TCC TCC CTC AAC CAC ACC ATC TTG GCT GGA GTT CAC AGC AAT GAC
R   S   A   S   S   L   N   H   T   I   L   A   G   V   H   S   N   D

               1089
CAC TCT GTG GTT TAG  3'
H   S   V   V   *
```

FIGURE 1B



Decoration 'Decoration #1': Box residues that match 80853 exactly.

FIGURE 2